H

hi

A transparent multi-backend proxy for Claude Code with hot-switching, cost tracking, and cross-platform setup.

About

hi is a multi-backend proxy for Claude Code that lets you switch between model providers without editing files by hand. It is built to be transparent, fast, and elastic, and it can launch a proxy with Claude Code or run as proxy-only when you do not need the agent attached. The app supports hot-switching between Claude, DeepSeek, MiniMax, GLM, Kimi, Qwen, and MIMO, and it restores changes on exit instead of leaving your setup modified.

What is hi

hi is a multi-backend proxy for Claude Code that sits between Claude Code and several

setup. It is designed to be transparent, so the proxy does not leave permanent file changes behind, and it restores its edits when you exit. As a Claude Code proxy, hi focuses on practical switching, simple setup, and supporting multiple providers from one tool.

The app is presented as a cross-platform single binary, with install scripts for Linux, macOS, and Windows. It also advertises cost tracking with persistence across sessions, which is useful if you want to keep an eye on usage while moving between backends. For users who work with Claude Code daily, hi is essentially a backend switcher and proxy manager in one.

Key features

- Multi-backend proxy for Claude Code with hot-switch commands.
- Supports switching among Claude, DeepSeek, MiniMax, GLM, Kimi, Qwen, and MIMO.
- Transparent behavior: no files are modified permanently, and settings are restored on exit.
- Cross-platform single binary distribution with OS-specific install commands.
- Cost tracking that persists across sessions.
- Auto-generated slash commands for quick backend changes.
- Two usage modes: `hi` to launch the proxy with Claude Code, or `hi proxy` for proxy-only operation.

A few model details are listed on the site as supported options, including Anthropic Claude Sonnet 4.6, Opus 4.8, and Haiku 4.5; DeepSeek V4 Pro and V4 Flash; MiniMax M3; Zhipu GLM 5.1; Moonshot Kimi K2.6; Alibaba Qwen 3.7-Max; and Xiaomi MIMO V2.5-Pro. The presence of model names in the documentation suggests that hi is meant to route requests across specific providers rather than act as a general AI chat app.

How to use it

- Install Claude Code first, because hi lists that as a prerequisite.
- Install hi using the script for your platform:
  - Linux/macOS: `curl -fsSL https://raw.githubusercontent.com/mars-base/hi/main/install.sh | sh`

Windows PowerShell: `irm https://raw.githubusercontent.com/mars-base/hi/main/install-windows.ps1 | iex`

Windows CMD: `curl -fsSL https://raw.githubusercontent.com/mars-base/hi/main/install-windows.cmd -o install.cmd && install.cmd && del install.cmd`

Run `hi init-config` to auto-detect settings, or set environment variables instead of using a config file.

If you want to edit settings manually, the config paths are `~/.hi/config.yaml` on Linux/macOS and `%USERPROFILE%\.hi\config.yaml` on Windows.

Start the proxy with `hi`, attach Claude Code to an existing proxy with `hi cc`, or run `hi proxy` if you only want proxy behavior.

Use slash commands like `/deepseek` or `/claude` to hot-switch backends when you need a different provider.

The setup flow is intentionally short, and the app supports both config-file and environment-variable workflows. The environment-variable examples shown on the site include `ANTHROPIC_API_KEY` and `DEEPSEEK_API_KEY`, which indicates that hi expects provider credentials to be supplied locally.

Who it is for

hi is for developers who already use Claude Code and want a cleaner way to move between backend providers. It is a good fit if you test different models, compare output quality, or need a proxy that can switch providers without manual reconfiguration.

It is also useful for users who care about operational details such as cost tracking and session persistence. Because hi supports multiple providers and emits slash commands for switching, it suits people who want a workflow tool rather than a consumer-facing chatbot.

Typical users include:

Claude Code power users who want backend flexibility.

Developers comparing Claude, DeepSeek, GLM, Qwen, Kimi, MiniMax, or MIMO.

People who want a proxy-only mode for controlled routing.

Users who prefer one binary and shell-based setup over a heavier client.

What to know before you use it

hi is not a substitute for Claude Code itself; the site explicitly says Claude Code must be installed first. It also depends on provider APIs and local keys, so it is not a zero-setup hosted service.

Important caveats:

- Model availability is limited to the providers and versions listed on the project site; if you need a different model, hi may not support it.
- The documentation does not describe pricing, so treat cost as whatever your provider accounts charge.
- Because it is a proxy, hi helps route and switch requests, but it does not replace the underlying model services.
- You should still consult the official provider docs for rate limits, model behavior, and account settings.
- The project is distributed under CC BY-NC 4.0 on GitHub, which is relevant if you intend to reuse or redistribute it.

In short, hi is a practical Claude Code proxy for backend switching, persistence, and cost tracking, but it works best when you already understand the provider accounts and API keys behind it. If you need the exact behavior of a model or the latest provider-side rules, use the official source for that backend rather than relying on the proxy alone.

FAQ

What is hi used for?

Is hi free to use?

Does hi modify my files permanently?

What do I need before installing hi?

How do I configure hi?

Can hi switch between different AI providers?

Is Claude Code required?

Does hi replace the model providers it connects to?

Sponsored

Belay – Awa... Featured
Developer Tools

Stays awake while your agents work.

Free # Developer Tools

View ↗

Partly Featured
Finance

Partly is a mobile app that lets friends quickly scan receipts, split any bill fairl...

Free # Finance

View ↗

Visit website ↗

hi

Pricing	Free
Category	AI
Platform	Linux, macOS, Windows, API
Launched	Aug 2026
Updated	yesterday
Verified	2 months ago

H hi

claude code # proxy # backend switching # multi-backend # ai tools # hot-switching # cli # cross-platform # open source # cost tracking

© 2026 Dir Hub

H hi